Investigation of the Relationship Between Susceptibility Loci for Hip Osteoarthritis and Dual X-Ray Absorptiometry–Derived Hip Shape in a Population-Based Cohort of Perimenopausal Women

Baird, Denis A; Paternoster, Lavinia; Gregory, Jennifer S; Faber, Benjamin G; Saunders, Fiona R; Giuraniuc, Claudiu V

Published in:
Arthritis and Rheumatology

DOI:
10.1002/art.40584

Publication date:
2018

Document Version
Peer reviewed version

DR. DENIS ANDREW BAIRD (Orcid ID : 0000-0003-4600-6013)

PROF. RICHARD M ASPDEN (Orcid ID : 0000-0002-0693-1194)

Article type : Full Length

Investigation of the relationship between susceptibility loci for hip osteoarthritis and DXA-derived hip shape in a population based cohort of peri-menopausal women

Denis A. Baird PhD[1], Lavinia Paternoster PhD[2], Jennifer S. Gregory PhD[3], Benjamin G. Faber BSc MBBS[1], Fiona R. Saunders PhD[3],Claudiu V. Giuraniuc PhD[3], Rebecca J. Barr PhD[4], Deborah A. Lawlor PhD[2], Richard M. Aspden PhD[3], Jonathan H. Tobias MD PhD[1]

[1]Musculoskeletal Research Unit, University of Bristol

[2]MRC Integrative Epidemiology Unit at the University of Bristol

[3]Arthitis and Musculoskeletal Medicine, University of Aberdeen

[4]Medicine Monitoring Unit, University of Dundee

Key words: Genetics, SSM, ALSPAC

Address for correspondence:-

Professor Jon Tobias, Musculoskeletal Research Unit, Learning and Research Building, Southmead Hospital, Bristol BS10 5NB

Jon.Tobias@bristol.ac.uk

Abstract

Objective: To investigate whether hip shape contributes to osteoarthritis (OA) development, we investigated relationships between known OA susceptibility loci and hip shape in a population-based cohort of peri-menopausal women.

Methods: Hip shape was measured using Statistical Shape Modelling, on hip DXA scans from mothers in the Avon Longitudinal Study of Parents and Children (ALSPAC). The proximal femur and superior acetabulum were outlined, and independent hip shape modes generated. In a sub-regional model, points were restricted to the acetabulum and superior femoral head. Associations between 11 OA-related SNPs, identified by literature search, and shape modes were analysed in a multivariate canonical correlation analysis.

Results: 3,111 females had genetic and hip shape data (mean 48 years). The *KLHDC5-PTHLH* rs10492367 OA risk allele was associated with a wider upper femur in the whole shape model ($P=1x10^{-5}$). The *DOT1L* rs12982744 OA risk allele was associated with reduced superior joint space in the sub-regional shape model ($P=2x10^{-3}$). *COL11A1* rs4907986 OA risk allele was associated with lateral displacement of the femoral head relative to the acetabulum in the sub-regional shape model ($P=5x10^{-4}$). Regional association plots identified an additional *COL11A1* locus in moderate LD with rs4907986, which was more strongly associated with hip shape (rs10047217, $P=6x10^{-6}$). Co-localisation analysis indicated sharing of genetic signals for hip shape and hip OA for the *KLHDC5-PTHLH* and *COL11A1* loci.

Conclusion: Hip OA susceptibility loci were associated with shape in this study suggesting that these loci (and potentially yet to identified hip OA loci) could contribute to hip OA in later life via perturbing biologic pathways that mediate morphology development.

Introduction

Genetic association studies have identified several susceptibility loci for hip osteoarthritis (OA), which may ultimately improve understanding of pathogenesis and lead to the development of new treatments (1). Whereas initial studies were based on case collections, defined by radiographic hip osteoarthritis (RHOA) or history of joint replacement, availability of more precise endophenotypes may increase the chances of genetic discovery, as suggested for hypertrophic OA and superior joint space narrowing (2). There is also considerable interest in the role of shape abnormalities in hip OA, exemplified by cam-type deformities, caused by extra bone growth around the anterolateral aspect of the femoral head leading to femoro-acetabular impingement (FAI) (3), which are associated with premature onset of OA (4, 5). In the Rotterdam study, individuals with cam deformity and acetabular dysplasia were reported to have around a 2 fold increased risk of radiographic hip OA compared to controls (6).

Rather than investigating specific geometric/anatomical measures, statistical shape modelling (SSM) can be employed, whereby Principal Component Analysis (PCA) is used to derive a set of orthogonal hip shape modes (HSMs), which together provide a more complete description of hip shape (7). Several SNPs have been associated with changes in hip shape using SSMs. An intronic *DIO2* SNP was associated with the first HSM in a small study comprising of 190 sib pairs and two trios of Dutch ancestry, using a model built on the medial femur, acetabulum and pelvis (8). Two independent missense *FRZB* SNPs were associated with the second HSM in a nested case control study of older women from the study of osteoporotic fractures (n=1,046) using a superior femur model (9). An intronic SNPs within *ASTN2* and a 3 prime UTR SNP within *GLT8D1* were associated with distinct HSMs

in 929 subjects with unilateral hip OA, in a model built on the superior femur from the unaffected hip. In the same study, a SNP located near *IFRD1* was also associated with a combination of three distinct modes in a multivariate canonical correlation analysis (10).

The above studies were based on older participants with established hip OA, making it difficult to distinguish genetic associations with hip shape representing part of the OA phenotype, from shape changes predisposing to subsequent OA development. To investigate the latter, genetic associations with hip shape need to be examined in younger people from the general population. Such analyses, however, are difficult to undertake since hip radiographs are generally avoided in asymptomatic younger individuals due to their relatively high radiation dose. An alternative approach is to use DXA-based measures, which are associated with considerably lower radiation exposure, and are available in several large population-based cohorts. For example, HSMs derived from hip DXA scans, based on an 85 point model comprising the femoral head, acetabulum and femoral neck were recently reported to be associated with radiographic OA in the Tasmanian Older Adult Cohort (11).

To establish whether genetic risk factors for hip OA partly act through alterations in shape, we examined associations between established hip OA susceptibility loci, and DXA-derived hip shape, in a population-based cohort of peri-menopausal women. Rather than performing a series of univariate tests on each of the hip shape modes, in order to model the effects of HSMs together we applied the multivariate canonical correlation approach (CCA) described by Lindner et al (10). In addition, we investigated the use of sub-regional HSMs, intended to focus on shape abnormalities involved in the development of hip OA.

Materials and Methods

Hip OA Candidate Genes The literature was examined to find studies which associated single nucleotide polymorphisms (SNPs) with hip OA. The keywords Osteoarthritis AND genome were initially searched for in Pubmed and the publications returned from this search were manually reviewed. Genetic associations with hip OA and overall OA (as indicated by total joint replacement) were considered only (i.e. associations with other OA site, for example, knee or hand OA were excluded). A significance threshold of $P < 5x10^{-8}$ was used to identify associations from GWASs, and a significance threshold of $P < 5x10^{-5}$ was applied to candidate gene studies. Since the present investigation was based on a female-only cohort, genetic associations reaching the threshold in males only (i.e. not in combined genders and females) were excluded.

Study participants

The study was based on mothers from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort, a longitudinal birth cohort that recruited pregnant women resident in a geographical area in the South West of England, UK, with an expected delivery date between April 1, 1991, and December 31, 1992. The present study uses data from follow-up research clinics of mothers between 2008 and 2011 (12, 13). All eligible mothers (*i.e.* still engaged with the study; alive with known contact details and who had not withdrawn their consent) were invited to these assessments. Of 11,264 (82%) women invited, 4,834 (43%) attended. Informed consent was collected for all in line with the Declaration of Helsinki (14). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees (North Somerset & South Bristol Research Ethics Committee: 08/H0106/96). The study website contains details of all the data that is available

through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/).

DXA protocol

Hip dual-energy X-ray absorptiometry (DXA) scans, performed by GE Lunar Prodigy Scanner (Madison, WI, USA), were acquired and analysed according to manufacturer's standard scanning and positioning protocol. At the same visit, height was measured to the nearest 0.1 cm using a Harpenden Stadiometer (Holtain Ltd. Crymych, UK) without shoes, and weight to the nearest 50 g using Tanita weighing scales (Tanita UK Ltd, Uxbridge).

Statistical shape modelling

Hip DXA scans were uploaded to the SHAPE software (University of Aberdeen) to conduct the SSM. SHAPE automatically placed points around anatomical landmarks of the upper femur and adjacent acetabulum according to a pre-defined 53 point template (see Figure 1). Repeatability was evaluated by comparing two experienced operators on a sample of 100 hips from ALSPAC kids (age 18). Results showed high consistency between both operators with the mean difference (across all 53 points on the template) being 1.8 pixels (SD=0.4, range = 1.0 to 2.9). All operators were trained on a separate set of 100 hips prior to the main analysis to ensure the same high standard of placement was maintained. Hip shape size and rotation were standardized by Procrustes analysis, and Principal Component Analysis (PCA) performed to produce a set of HSMs describing linearly independent variations in shape present within the entire sample (15, 16). To aid comparison with other studies, we used HSMs previously generated in ALSPAC from a larger PCA combining scans from multiple cohorts, for the purposes of performing a GWAS meta-analysis (17). The mode scores follow

a standard Normal distribution and describe SD changes from the mean shape. As a quality control measure, outliers (mode scores above or below 4SDs, n=48) were manually checked by two operators, and point placement corrected where necessary to ensure points were positioned accurately on the bone edge. The first ten modes, accounting for 85% of total variance, were taken forward for genetic association studies (see Supplementary Figure 1). Three further sub-regional models were built, based on a subset of points, of which the first five modes (explaining > 85% variability within each sub-model) were used for genetic association analyses: femoral head (Supplementary Figure 2), cam-type deformity (Supplementary Figure 3) and superior joint space (Supplementary Figure 4).

Genotypic data

Genome-wide SNP data available on the ALSPAC Mother (n=10,015) samples were used for this study. SNPs with a genotype missingness rate greater than 5%, extreme deviation from Hardy-Weinberg equilibrium ($P < 5x10^{-7}$), and a minor allele frequency (MAF) of less than 1% were removed. Individuals with greater than 5% missingness, extreme heterozygosity, evidence of population stratification and cryptic relatedness (IBD > 0.125) were excluded. Details of this pipeline are reported elsewhere (18). Genotypic data were phased using ShapeIT version 2 and imputed on IMPUTE version 2.2.2 using the Haplotype Reference Consortium (HRC) version 1 panel. SNPs with an imputation info score of less than 0.8 and MAF < 0.01 were subsequently removed, and genotypes were derived from the imputed data using GTOOL and a hard genotype call threshold of 0.9. The genotypes of all the candidate OA SNPs identified were then extracted for the association analysis. The surrounding genomic region was also extracted in order to generate the regional association plots and conduct the colocalization analysis.

Statistical analysis

Canonical Correlation Analysis (CCA) was used to simultaneously test for association between a given SNP and multiple shape modes, using the canon command in Stata version 13.1, which assumes an allele dose genetic model. The first ten ancestry principal components and age were regressed on each of the HSMs to obtain residual scores, which were then used in the CCA models. As a sensitivity analysis, SNP associations were also pre-adjusted for height before conducting the CCA. A linear combination of the first ten HSMs for the whole shape model and the first five modes in the sub-regional models were tested in the CCA. To visualize shape effects, plots showing the deviation from the mean shape according to the risk allele for hip OA, based on the significant modes identified from CCA, were generated in SHAPE. Where CCA revealed evidence of associations with multiple modes, the canonical weightings of the different modes were used to plot the composite shape change observed for the SNP. A Bonferroni-adjusted -p-value was applied of $P<1.25x10^{-3}$, to account for the multiple SNPs tested. To conduct conditional analysis on the multiple SNP signals detected within *COL11A1,* a MANOVA regression framework was employed in Stata using the mvreg command.

Approximate Bayesian co-localization analysis was carried to determine the probability of hip shape and hip OA traits sharing the same casual SNP within the genomic regions identified. This method calculates a posterior probability for five separate hypotheses: H_0: SNP is associated with neither trait, H_1: SNP is associated with hip shape only, H_2: SNP is associated with hip OA only, H_3: both traits are associated but with independent causal SNPs and H_4: both traits are associated with the same causal SNP. High posterior probability for H_4 in comparison to the other four hypotheses is evidence for colocalization of genetic signal.

The coloc.abf function in the coloc package in R was used (19-21). Three prior probabilities were required to be set: the expected proportion of SNPs associated with hip shape (p_1), with hip OA (p_2) and with both these traits (p_{12}). The default priors used in the coloc program was used which were $p_1=1x10^{-4}$, $p_2 = 1x10^{-4}$ and $p_{12} = 1x10^{-6}$. CCA p-values and MAF obtained from this study were used as the input for the hip shape associations and p-values and MAF from the arcOGEN GWAS (discovery cohort) (22) were used for the hip OA associations.

Results

OA Candidate genes

11 SNPs were selected from the literature as having evidence for association with hip OA (Table 1). In the majority of the studies, OA cases comprised of RHOA, severe symptomatic hip OA, total hip replacement (THR), or a combination.

Associations between OA candidate genes and hip shape

3,111 ALSPAC mothers had valid genetic and DXA hip shape data. At the time of DXA scans these women were mean age 48 years (range 34-61 years), 2,191 (70.4%) women were pre-menopausal and 920 (29.6%) 920 (29.6%) post-menopausal. All hip shape modes were Normally distributed. Results of genetic association analysis based on the hip shape models are shown in Table 2. *KLHDC5-PTHLH* rs10492367 SNP showed a Bonferroni significant association (ρ (canonical correlation coefficient) = 0.11; $P=1x10^{-5}$), whereas weaker evidence of association (nominal significance, $P<0.05$) was observed for *COL11A1* rs4907986. For the superior joint space model, a Bonferroni significant association was observed for *COL11A1* rs4907986 ($\rho=0.085$; $P=5x10^{-4}$), and a suggestive association for *DOT1L* rs12982744 ($\rho=0.077$; $P=2x10^{-3}$) which was on the borderline of Bonferroni significance, whereas

KLHDC5-PTHLH rs10492367 (P=0.02) showed weaker evidence of association (Table 2). Only relatively weak evidence of associations (P<0.05) were observed for the femoral head shape model (*KLHDC5-PTHLH* rs1049236 and *COL11A1* rs4907986), no nominally significant associations were observed for the cam-type deformity model (Table 2). None of the associations investigated attenuated after additional adjustment for height (Supplementary Table 1).

***KLHDC5-PTHLH* rs10492367**

The relationship between *KLHDC5-PTHLH* rs10492367 and whole hip shape was predominantly explained by a positive association with mode 1, indicating greater width to height ratio. A positive association with mode 10 was also detected by CCA (Supplementary Table 2). Subsequently, weightings (x2 to magnify effects for illustration) from CCA were entered into SHAPE. Modelling both modes together using CCA indicated that the *KLHDC5-PTHLH* rs10492367 risk allele is associated with a relatively wide and short upper femur (Figure 2A). Regional association plot for the genetic association with whole hip shape from the multivariate analysis showed similar associations for several neighbouring SNPs in high linkage disequilibrium (LD) with rs10492367 (Figure 2B). Further regional association plots comparing hip shape associations from this study with hip OA associations from arcoGEN (22) revealed a similar association pattern for each phenotype. Co-localization analysis confirmed that both these traits shared a single causal SNP (Posterior Probability (PP) = 99%).

***DOT1L* rs12982744**

The relationship between *DOT1L* rs12982744 and shape in superior joint space model was primarily explained by a positive association with mode five, indicating reduced joint space.

A positive association with mode 1 also contributed to this relationship, indicating medial displacement of the femoral head relative to the acetabulum (Supplementary Table 3). Modelling both modes simultaneously using CCA suggested that the *DOT1L* rs12982744 risk allele is associated with reduced supero-lateral joint space (Figure 3A). In the regional association plot, several neighbouring SNPs in moderately high LD with *DOT1L* rs12982744 showed equivalent shape associations (Figure 3B). Further regional association plots comparing shape associations with hip OA associations from arcOGEN (22) suggested these signals may overlap (inference from formal colocalization analysis was limited by the relatively weak associations of this locus with hip shape and hip OA).

COL11A1 rs4907986

The relationships between the rs4907986 locus and shape in the superior joint space model was largely explained by an inverse association with mode 1, reflecting lateral displacement of the femoral head relative to the acetabulum. In addition, an inverse association was observed with mode 5, indicating wider joint space (Supplementary Table 3). Modelling both these modes simultaneously using CCA suggested that the risk allele is associated with lateral displacement of the femoral head combined with widening of the supero-lateral joint space (Figure 4A). Regional association plots of this multivariate association showed a substantial number of signals in moderately strong linkage across the *COL11A1* gene. An additional association with superior joint space shape of a similar strength was detected with rs2615977 (P=0.00048), which was previously identified as associated with hip OA in the stage one arcOGEN GWAS (OR = 1.10, $P=1.10x10^{-5}$ combined discovery and replication) (23). rs2615977 and rs4907986 occurred in two distinct LD blocks and were in moderate LD with each other ($r^2 = 0.353$) (Figure 4B). A further SNP (rs10047217) was identified which had a stronger association with shape ($P=3x10^{-6}$) than either of these two SNPs; after conditioning

on rs10047217 there was no evidence that either rs2615977 or rs4907986 were independently associated with shape (P=0.68 and P=0.13 respectively). Regional association plots comparing shape associations with hip OA associations from arcOGEN (22) indicated sharing of association signals between each of the traits. Co-localization analysis suggested that both these traits shared a single causal SNP (PP= 70%).

Discussion

We found that several known OA susceptibility SNPs are associated with hip shape in whole and/or sub-regional hip models, in a cohort of peri menopausal women. Specifically, *KLHDC5-PTHLH* rs10492367 was associated with a wider upper femur, *DOTL1* rs12982744 with reduced supero-lateral joint space, and *COL11A1* rs2615977 and rs4907986 with lateral displacement of the femoral head relative to the acetabulum. Several shape differences have previously been reported to be associated with OA, but it is often difficult to distinguish changes causing OA from those secondary to it. By studying individuals at an age when they are likely to be disease free, indices of hip shape are likely to represent causal pathways in OA development as opposed to changes consequent to OA. Co-localisation analysis supported the conclusion of single causal variants giving rise to both hip shape and OA associations (rather than the pattern being driven by horizontal pleiotropy as a result of linkage disequilibrium) for the *KLHDC5-PTHLH* and *COL11A1* loci. However, our analysis could not rule out alternative mechanisms of pleiotropy and so are not confirmative of a causal mechanism.

KLHDC5-PTHLH rs10492367 had previously been found to be associated with hip OA following a GWAS with replication in the arcOGEN case control study (22). This locus is downstream to the gene encoding parathyroid hormone-related protein (PTHrp), an important

regulator of endochondral bone development (24). Conceivably, the difference in height to width ratio associated with *KLHDC5-PTHLH* rs10492367 is a consequence of altered endochondral bone development. However, it is unclear how this shape change might cause hip OA. Previous DXA-based studies have reported associations between altered femoral geometry and risk of subsequent hip OA, exemplified by a report that wider femoral neck and medial displacement of the centroid position are risk factors for incident radiographic hip OA (25). However, we are not aware of any previous studies that directly related height to width ratio of the upper femur to risk of hip OA.

DOT1L rs12982744, which we found to be associated with supero-lateral joint space, was previously reported to be associated with hip OA in a look-up of the TREAT-OA consortium (26). This SNP had previously been identified from a GWAS of joint space width, assessed using hip radiographs from older individuals (predominantly in their mid-sixties) from population based cohorts (27). This apparent effect on cartilage integrity is consistent with observations that DOT-1-like histone H3 methyltransferase, which is encoded by *DOT1L,* influences chondrogenesis (27). That *DOT1L* rs12982744 is also associated with joint space in individuals in their late forties as studied here, suggests *DOT1L* variation affects joint space prior to the onset of symptomatic OA, consistent with a role in OA pathogenesis. As well as joint space narrowing, simultaneous modelling of both modes related to *DOT1L* rs12982744 suggested that the OA risk allele causes a progressive reduction in joint space laterally (see Figure 2). This appearance is somewhat reminiscent of FAI, and raises the possibility that bony changes also contribute to the increased hip OA risk associated with *DOT1L* rs12982744.

Of the two independent *COL11A1* candidate SNPs examined, rs4907986, which had previously been associated with hip OA in a meta-analysis of OA candidate genes (28), showed the strongest evidence for association in the superior joint space narrowing model. However, further analysis on the *COL11A1* region did reveal that an alternative SNP (rs10047217), which was in moderate LD with rs4907986, had a stronger association with hip shape, suggesting that this SNP could be the causal SNP. *COL11A1* encodes COL11A1, a fibrillary protein within cartilage, mutations of which are known to cause severe osteoarthritis secondary to spondylo-epiphyseal dysplasia as in Stickler's syndrome (29). Our observation that *COL11A1* risk alleles are associated with lateral displacement of the femoral head relative to the acetabulum is reminiscent of acetabular dysplasia. The latter is associated with a reduction in acetabular coverage of the femoral head which is thought to increase hip OA risk through increased contact stress arising from the smaller weight-bearing surfaces. Acetabular coverage can be quantified on hip radiographs through measurement of lateral centre-edge angle, a reduction of which has been associated with increased risk of incident hip OA (30, 31) and has also been associated with an increased risk of total hip replacement, independently of radiographic OA (32).

There was some overlap in findings across different models. For example, *KLHDC5-PTHLH* rs10492367 showed weak associations in femoral head and acetabular models, and *DOT1L* rs12982744 and *COL11A1* rs4907986 weak association in the whole shape model, which is to be expected given the overlap in points between models. However, that *DOT1L* and *COL11A1* SNPs showed considerably stronger associations in the acetabular model suggests the use of sub-regional models is helpful in detecting genetic associations with hip shape, presumably reflecting the localised nature of underlying biological processes. Whereas several other genetic susceptibility SNPs for hip OA have previously been reported in

association with SSM-derived hip shape (10, 33, 34), we did not find these to be related to DXA-derived hip shape. A possible explanation for this is that the latter studies were based on hip OA cases, which may have resulted in the detection of SNPs associated with features related to established OA, as opposed to factors involved in initial pathogenesis.

Strengths and limitations

To our knowledge, this is the first study to describe genetic associations with DXA-derived hip shape. A further strength is our use of sub-regional models, which enhanced discovery of genetic associations. Although the modes in the sub-regional models will be partly correlated with certain modes in the whole shape model, as the sub-regional models describe a more precise phenotype they will be better powered to detect SNPs associated with pathways that mediate localised changes to hip morphology. Another strength was the relatively young population based cohort, enabling us to identify genetic associations with features related to OA development rather than established OA, made possible by our use of DXA images which are more widely available in population-based cohorts compared with radiographs. One of the main limitations is that the present study was restricted to females, and it is unclear whether equivalent associations exist in males. The relatively low resolution of DXA compared with radiographs represents a significant disadvantage in detecting shape changes, not the least due to the difficulty of detecting and excluding artefacts such as osteophytes. However, the consistency of our findings for *DOT1L* rs12982744, which was associated with joint space narrowing both in our study and a previous radiographic study albeit in older individuals (27), supports the fact that DXA-derived hip shape provides sufficient accuracy for genetic association studies. In addition, although we have interpreted certain hip shape modes in terms of specific shape alterations, since each mode encompasses multiple shape changes, it is difficult to be certain of which component is relevant for pathology. Therefore,

although our results suggested genetic associations with features such as FAI and acetabular un-coverage, this requires confirmation by relating SSM-derived shape to measures such as lateral centre-edge angle used to define these abnormalities on radiographs.

Conclusions

Having examined 11 SNPs previously reported to be associated with risk of hip OA, three SNPs (in three loci) were found to be associated with DXA-derived hip shape, one in the whole shape model, and two in a novel sub-regional model restricted to the upper femoral head and acetabulum. Shape changes found to be related to OA risk alleles in the sub-regional model were reminiscent of FAI and acetabular dysplasia, which have previously been implicated in OA pathogenesis. In contrast, in the whole shape model, the OA risk allele was related to upper femur width to height ratio, pointing to a potentially novel disease mechanism. The more precise understanding of hip OA pathogenesis gained through studying SSM endophenotypes may enable improved targeting of functional studies leading to better understanding of molecular mechanisms and potential improved resolution to determine causal SNPs from GWASs. Our findings need replication, but they suggest that hip shape may be a marker of future OA risk and could potentially provide a means of early identification and prevention. As hip OA is a complex phenotype, loci are likely to be associated with different pathways to the disease. We infer that the hip OA SNPs identified in this hip shape study will, at least in part, act through impacting molecular pathways which mediate femur morphology development.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We would also like to thank the University of Aberdeen Data Management Team for software support. This study was funded by Arthritis Research UK (20244), which provided salary funding for DB, Wellcome Trust (WT092830M) and UK Medical Research Council (G1001357), which provided funds for collection of hip shape, and the Wellcome Trust (WT088806), which funded the genotyping. LP and DAL work in a unit that recieves support from the University of Bristol and the UK Medical Research Council (MC_ UU_ 12013/4 & MC_UU_12013/5). BF was supported by an Elizabeth Blackwell Institute Clinical Research Primer Scheme. The UK Medical Research Council and the Wellcome Trust (ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and does not necessarily reflect the views of any funders. None of the funders had any influence on data collection, analysis, interpretation of the results or writing of the paper. DB will serve as the guarantor of the paper.

Figure Legends

Figure 1

Points used in different hip shape models. Hip shape template annotated with the points included in the (a) whole shape model (points 0,1,44, 45, and 57 (marked with x) were excluded from the final model due to inconsistency of point placement); (b) superior acetabular model, comprising points 18 to 23 on the femur, and points 56 to 51 on the acetabulum; (c) femoral head model, comprising points 9 to 23 on the femoral head; and (d)

the cam-type deformity model, comprising points 21 to 29 on the femoral head and neck. Key points, representing anatomical reference points, are shown in red.

Figure 2

(A) Effect of the rs10492367 *KLHDC5-PTHLH* locus on whole hip shape, modelled according to the hip OA risk allele (T allele), using loadings from CCA correlation. Canonical function coefficients (x2) were entered into SHAPE, for those modes showing significant associations (see Supplementary Table 3). Mean shape (solid line), composite shape change observed (dotted line). (B) Regional association plot showing the multivariate associations (-log10 P-value) for the whole hip model for SNPs surrounding *KLHDC5-PTHLH* rs104292367 locus (annotated as purple diamond). SNP associations are coloured according to strength of LD with rs104292367.

Figure 3

(A) Effect of the rs12982744 *DOT1L* locus on hip shape in the acetabular model, modelled according to hip OA risk allele (C allele), using loadings from CCA correlation. Canonical function coefficients (x2) were entered into SHAPE, for those modes showing significant associations (see Supplementary Table 4). Mean shape (solid line), composite shape change observed (dotted line). (B) Regional association plot showing the multivariate associations for the acetabular model for SNPs within *DOT1L*. SNP associations are coloured according to strength of LD with the index SNP (rs12982744, purple diamond).

Figure 4

(A) Effects of rs4907986 *COL11A1* locus on hip shape in the acetabular model, modelled according to the hip OA risk allele (rs4907986: T allele), using loadings from CCA correlation. Canonical function coefficients (x2) were entered into SHAPE, for those modes showing significant associations (see Supplementary Table 4). Mean shape (solid line), composite shape change observed (dotted line). (B) Regional association plot showing the multivariate associations for the acetabular model for SNPs within *COL11A1*. SNP

associations are coloured according to LD with the index SNP (rs4907986, purple diamond). Rs2615977 SNP which was also associated with hip OA in another study and the rs2615977 which was the top hit for hip shape in this study are also highlighted with arrows in the plot.

References

1. Warner SC, Valdes AM. Genetic association studies in osteoarthritis: is it fairytale? Curr Opin Rheumatol. 2017;29(1):103-9.

2. Panoutsopoulou K, Thiagarajah S, Zengini E, Day-Williams AG, Ramos YF, Meessen JM, et al. Radiographic endophenotyping in hip osteoarthritis improves the precision of genetic association analysis. Ann Rheum Dis. 2017;76(7):1199-206.

3. Ganz R, Parvizi J, Beck M, Leunig M, Nötzli H, Siebenrock KA. Femoroacetabular Impingement: A Cause for Osteoarthritis of the Hip. Clinical Orthopaedics and Related Research. 2003;417:112-20.

4. Beck M, Kalhor M, Leunig M, Ganz R. Hip morphology influences the pattern of damage to the acetabular cartilage: femoroacetabular impingement as a cause of early osteoarthritis of the hip. J Bone Joint Surg Br. 2005;87(7):1012-8.

5. Murray RO. The aetiology of primary osteoarthritis of the hip. Br J Radiol. 1965;38(455):810-24.

6. Saberi Hosnijeh F, Zuiderwijk ME, Versteeg M, Smeele HT, Hofman A, Uitterlinden AG, et al. Cam Deformity and Acetabular Dysplasia as Risk Factors for Hip Osteoarthritis. Arthritis Rheumatol. 2017;69(1):86-93.

7. Gregory JS, Waarsing JH, Day J, Pols HA, Reijman M, Weinans H, et al. Early identification of radiographic osteoarthritis of the hip using an active shape model to quantify changes in bone morphometric features: Can hip shape tell us anything about the progression of osteoarthritis? Arthritis & Rheumatism. 2007;56(11):3634-43.

8. Waarsing JH, Kloppenburg M, Slagboom PE, Kroon HM, Houwing-Duistermaat JJ, Weinans H, et al. Osteoarthritis susceptibility genes influence the association between hip morphology and osteoarthritis. Arthritis Rheum. 2011;63(5):1349-54.

9. Baker-Lepain JC, Lynch JA, Parimi N, McCulloch CE, Nevitt MC, Corr M, et al. Variant alleles of the Wnt antagonist FRZB are determinants of hip shape and modify the relationship between hip shape and osteoarthritis. Arthritis Rheum. 2012;64(5):1457-65.

10. Lindner C, Thiagarajah S, Wilkinson JM, Panoutsopoulou K, Day-Williams AG, arc OC, et al. Investigation of association between hip osteoarthritis susceptibility loci and radiographic proximal femur shape. Arthritis Rheumatol. 2015;67(8):2076-84.

11. Ahedi HG, Aspden RM, Blizzard LC, Saunders FR, Cicuttini FM, Aitken DA, et al. Hip shape as a predictor of osteoarthritis progression in a prospective population cohort. Arthritis Care Res (Hoboken). 2016.

12. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol. 2013;42(1):111-27.

13. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. Int J Epidemiol. 2013;42(1):97-110.

14. Rickham P. Human experimentation. Code of Ethics of the World Medical Association. Declaration of Helsinki. British Medical Journal. 1964;2(5402):177.

15. Gregory JS, Aspden RM. Femoral geometry as a risk factor for osteoporotic hip fracture in men and women. Med Eng Phys. 2008;30(10):1275-86.

16. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models - Their Training and Application. Comput Vis Image Und. 1995;61(1):38-59.

17. Baird D, Evans D, Gregory JS, Aspden RM, Kinyua F, Spector T, et al. Adult hip shape is influenced by variation in genes involved in endochondral bone formation: findings from a genome-wide association study followed by meta-analysis. Journal of Bone and Mineral Research. 2017;32:S175.

18. Evans DM, Zhu G, Dy V, Heath AC, Madden PA, Kemp JP, et al. Genome-wide association study identifies loci affecting blood copper, selenium and zinc. Hum Mol Genet. 2013;22(19):3998-4006.

19. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10(5):e1004383.

20. Plagnol V, Smyth DJ, Todd JA, Clayton DG. Statistical independence of the colocalized association signals for type 1 diabetes and RPS26 gene expression on chromosome 12q13. Biostatistics. 2009;10(2):327-34.

21. Wallace C, Rotival M, Cooper JD, Rice CM, Yang JH, McNeill M, et al. Statistical colocalization of monocyte gene expression and genetic risk variants for type 1 diabetes. Hum Mol Genet. 2012;21(12):2815-24.

22. Zeggini E, Panoutsopoulou K, Southam L, Rayner NW, Day-Williams AG, Lopes MC, et al. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. Lancet. 2012;380(9844):815-23.

23. Panoutsopoulou K, Southam L, Elliott KS, Wrayner N, Zhai G, Beazley C, et al. Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. Ann Rheum Dis. 2011;70(5):864-7.

24. Lanske B, Karaplis AC, Lee K, Luz A, Vortkamp A, Pirro A, et al. PTH/PTHrP receptor in early development and Indian hedgehog-regulated bone growth. Science. 1996;273(5275):663-6.

25. Javaid MK, Lane NE, Mackey DC, Lui LY, Arden NK, Beck TJ, et al. Changes in proximal femoral mineral geometry precede the onset of radiographic hip osteoarthritis: The study of osteoporotic fractures. Arthritis Rheum. 2009;60(7):2028-36.

26. Evangelou E, Valdes AM, Castano-Betancourt MC, Doherty M, Doherty S, Esko T, et al. The DOT1L rs12982744 polymorphism is associated with osteoarthritis of the hip with genome-wide statistical significance in males. Ann Rheum Dis. 2013;72(7):1264-5.

27. Castano Betancourt MC, Cailotto F, Kerkhof HJ, Cornelis FM, Doherty SA, Hart DJ, et al. Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. Proc Natl Acad Sci U S A. 2012;109(21):8218-23.

28. Rodriguez-Fontenla C, Calaza M, Evangelou E, Valdes AM, Arden N, Blanco FJ, et al. Assessment of osteoarthritis candidate genes in a meta-analysis of nine genome-wide association studies. Arthritis Rheumatol. 2014;66(4):940-9.

29. Acke FR, Malfait F, Vanakker OM, Steyaert W, De Leeneer K, Mortier G, et al. Novel pathogenic COL11A1/COL11A2 variants in Stickler syndrome detected by targeted NGS and exome sequencing. Mol Genet Metab. 2014;113(3):230-5.

30. Lane NE, Lin P, Christiansen L, Gore LR, Williams EN, Hochberg MC, et al. Association of mild acetabular dysplasia with an increased risk of incident hip osteoarthritis in elderly white women: the study of osteoporotic fractures. Arthritis Rheum. 2000;43(2):400-4.

31. Reijman M, Hazes JM, Pols HA, Koes BW, Bierma-Zeinstra SM. Acetabular dysplasia predicts incident osteoarthritis of the hip: the Rotterdam study. Arthritis Rheum. 2005;52(3):787-93.

32. Barr RJ, Gregory JS, Reid DM, Aspden RM, Yoshida K, Hosie G, et al. Predicting OA progression to total hip replacement: can we do better than risk factors alone using active shape modelling as an imaging biomarker? Rheumatology (Oxford). 2012;51(3):562-70.

33. Baker-LePain JC, Lane NE. Relationship between joint shape and the development of osteoarthritis. Curr Opin Rheumatol. 2010;22(5):538-43.

34. Waarsing JH, Rozendaal RM, Verhaar JA, Bierma-Zeinstra SM, Weinans H. A statistical model of shape and density of the proximal femur in relation to radiological and clinical OA of the hip. Osteoarthritis Cartilage. 2010;18(6):787-94.

35. Evangelou E, Kerkhof HJ, Styrkarsdottir U, Ntzani EE, Bos SD, Esko T, et al. A meta-analysis of genome-wide association studies identifies novel variants associated with osteoarthritis of the hip. Ann Rheum Dis. 2014;73(12):2130-6.

36. Day-Williams Aaron G, Southam L, Panoutsopoulou K, Rayner Nigel W, Esko T, Estrada K, et al. A Variant in MCF2L Is Associated with Osteoarthritis. American Journal of Human Genetics. 2011;89(3):446-50.

37. Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S, et al. A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. Nat Genet. 2007;39(4):529-33.

38. Rodriguez-Fontenla C, Calaza M, Evangelou E, Valdes AM, Arden N, Blanco FJ, et al. Assessment of Osteoarthritis Candidate Genes in a Meta-Analysis of Nine Genome-Wide Association Studies. Arthritis Rheumatol. 2014;66(4):940-9.

Table 1 Hip OA candidate SNPs selected in this study from (a) GWASs (b) candidate gene studies.

Study type	SNP	EA	NON EA	gene	phenotypic association	OR	P	Reference
GWAS	rs6976[1]	T	C	*GLT8D1*	TJR	1.12	$7.24x10^{-11}$	(22)
	rs4836732[2]	C	T	*ASTN2*	THR (females)	1.2	$6.11x10^{-10}$	(22)
	rs6094710	A	G	*NCOA3*	RHOA	1.28	$7.90x10^{-9}$	(35)
	rs10492367	T	G	*KLHDC5-PTHLH*	RHOA/THR	1.14	$1.48x10^{-8}$	(22)
	rs835487	G	A	*CHST11*	THR	1.13	$1.64x10^{-8}$	(22)
	rs11842874	G	A	*MCF2L*	ROA/TJR	0.85	$2.1x10^{-8}$	(36)
Candidate Gene	rs143383	C	T	*GDF5*	RHOA (Asians)	1.79	$1.8x10^{-13}$	(37)
	rs12982744	G	C	*DOT1L*	THR/RHOA	0.91	$8.1x10^{-8}$	(35)
	rs4907986	C	T	*COL11A1*	RHOA/THR	0.89	$1.29x10^{-5}$	(38)
	rs1241164	T	C	*COL11A1*	RHOA/THR	0.82	$1.47x10^{-5}$	(38)
	rs2862851	T	C	*TGFA*	RHOA/THR	0.059	$4.3x10^{-5}$	(27)

The final combined statistic from aggregating the effects across all cohorts involved in the meta-analysis is reported (for all studies except rs143383 where a meta-analysis was not conducted). Significance threshold to include SNPs from GWASs was $P<5x10^{-8}$ and from candidate gene studies was $P < 5x10^{-5}$. OA phenotypes comprised RHOA (radiographic hip OA), JSN (joint space narrowing), TJR (total joint replacement) and THR (total hip replacement). EA allele coded as the minor allele in the sample. Pairwise r^2 between rs4907986 and rs1241164 = 0.081.

Table 2 Multivariate associations for the hip OA candidate gene loci across the whole hip and different sub-regional shape models.

		whole hip		superior JSN		femoral head		cam-type deformity	
SNP	locus	ρ	P	ρ	P	ρ	P	ρ	P
rs10492367	*KLHDC5-PTHLH*	0.11	**0.000014**	0.064	0.024*	0.064	0.024*	0.024	0.87
rs4907986	*COL11A1*	0.078	0.046*	0.085	**0.00049**	0.061	0.044*	0.037	0.51
rs12982744	*DOT1L*	0.074	0.07	0.077	0.0024*	0.05	0.17	0.043	0.34
rs6094710	*NCOA3*	0.074	0.078	0.038	0.5	0.016	0.98	0.053	0.12
rs143383	*GDF5*	0.072	0.11	0.04	0.45	0.036	0.57	0.027	0.81
rs4836732	*ASTN2*	0.068	0.15	0.053	0.12	0.019	0.95	0.032	0.68
rs6976	*GLT8D1*	0.053	0.57	0.042	0.36	0.039	0.44	0.019	0.96
rs835487	*CHST11*	0.052	0.59	0.049	0.18	0.042	0.35	0.042	0.36
rs11842874	*MCF2L*	0.05	0.64	0.032	0.66	0.043	0.34	0.029	0.76
rs1241164	*COL11A1*	0.049	0.68	0.052	0.13	0.02	0.93	0.032	0.69
rs2862851	*TGFA*	0.049	0.70	0.048	0.22	0.017	0.97	0.033	0.65

Canonical correlation (ρ) between genotype and hip shape in multivariate analysis. A Bonferroni corrected threshold of $P < 0.00125$ was used to identify associations (marked in bold); associations reaching nominal significance ($P < 0.05$) are starred.